

Mineração de Opinião aplicada a postagens do Twitter sobre o Ensino Remoto Emergencial em Institutos Federais

Opinion Mining applied to Twitter posts about Emergency Remote Teaching in Federal Institutes

Mateus F. L. Souza
Instituto Federal de Minas
Gerais
CEP: 34590-390, Sabará,
MG, Brasil
++55 31 2102-9370
mateusfilipe557
@gmail.com

Cristiane N. Targa
Instituto Federal de Minas
Gerais
CEP: 34590-390, Sabará,
MG, Brasil
++55 31 2102-9370
cristiane.targa
@ifmg.edu.br

Carlos A. Silva
Instituto Federal de Minas
Gerais
CEP: 34590-390, Sabará,
MG, Brasil
++55 31 2102-9370
carlos.silva
@ifmg.edu.br

ABSTRACT

In this work, is proposed on sentiment analysis of the academic community of Federal Institutes during Emergency Remote Teaching from 2020 to 2021, using opinion mining techniques from posts on the social network Twitter. Tweets were collected at five different times. The tweets were classified as positive, negative and neutral and the results obtained showed a prevalence of the neutrality of the feelings of the academic community, and a certain stabilization of the positive feelings from February 2021.

CCS Concepts

•Computing methodologies → Information extraction;

Keywords

Sentiment analysis; social networks; teaching; pandemic

RESUMO

Neste trabalho é proposto um estudo sobre a percepção e os sentimentos da comunidade acadêmica de Institutos Federais no período compreendendo o Ensino Remoto Emergencial de 2020 a 2021, utilizando técnicas de mineração de opinião a partir de postagens da rede social *Twitter*. Foram realizadas coletas de *tweets* em cinco momentos distintos, a fim de construir uma base de dados para a aplicação do método Multinomial Naive Bayes, no intuito de categorizar os sentimentos percebidos. Os *tweets* foram classificados em positivos, negativos e neutros e pelos resultados obtidos

percebeu-se uma prevalência da neutralidade dos sentimentos da comunidade acadêmica, além de uma certa estabilização dos sentimentos mais positivos a partir de fevereiro de 2021.

Palavras-chave

Análise de sentimentos; redes sociais; ensino; pandemia

1. INTRODUÇÃO

A pandemia do coronavírus, causada pelo vírus SARS-CoV-2, impactou toda a sociedade mundial, a partir de 2020. Diversos setores da sociedade foram afetados, tendo suas atividades reduzidas ou paralisadas, parcialmente ou totalmente. O setor educacional foi um dos mais atingidos neste período. Cerca de 1,6 bilhão de alunos de mais de 190 países ficaram longe das salas de aula em 2020, e no Brasil, dos 56 milhões de alunos matriculados na educação básica e superior, cerca de 20 milhões de alunos tiveram as aulas suspensas [8]. Os Institutos Federais de Educação, Ciência e Tecnologia, tiveram suas aulas presenciais suspensas a partir de março de 2020 [4].

Para amenizar os prejuízos escolares causados pelo novo coronavírus, o Ministério da Educação (MEC) autorizou a substituição de disciplinas presenciais por aulas que utilizassem meios e tecnologias de informação e comunicação em cursos que estavam em andamento (PORTARIA Nº 343, DE 17 DE MARÇO DE 2020¹). De acordo com [8], aproximadamente 32 milhões de alunos no Brasil passaram a ter aulas remotas. Como alternativa à paralisação das atividades presenciais de ensino, diversas instituições de ensino, entre elas os Institutos Federais, adotaram o Ensino Remoto Emergencial (ERE) como mecanismo de manutenção das atividades pedagógicas. Por este modelo, pressupunha o uso de tecnologias para a realização das atividades do ensino remoto.

As redes sociais já faziam parte da vida das pessoas e,

¹<https://www.in.gov.br/en/web/dou/-/portaria-n-343-de-17-de-marco-de-2020-248564376>

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

neste momento de quarentena, tornaram-se espaços ainda mais importantes de relacionamento e comunicação. Houve um crescimento considerável do uso das redes sociais, sobretudo buscando um meio de interação para suprir o contato físico e pessoal. Segundo o Sindicato Nacional das Empresas de Telefonia e de Serviço Móvel Celular e Pessoal², o tráfego da internet durante a pandemia aumentou 30%, com *home office*, aulas *online*, entretenimento e redes sociais. Diante deste contexto, propomos um estudo sobre a percepção e os sentimentos da comunidade acadêmica dos Institutos Federais por meio da rede social *Twitter*, a fim de identificar e classificar as opiniões da seara acadêmica durante o período do ERE de 2020 à 2021.

A principal motivação para sustentar a pesquisa, reside na importância que o tema possui para a sociedade atual. Com a pandemia do novo coronavírus, o ERE foi adotado em todos os níveis de ensino tanto em escolas públicas quanto privadas. O mesmo ocorreu com os Institutos Federais, os quais também adotaram o ERE em diversos campi. É de suma importância conhecer a opinião da comunidade acadêmica, a fim de identificar e possibilitar a minimização dos problemas e ajustes no sistema de ensino adotado.

Para o desenvolvimento deste trabalho utilizou-se como base de dados *tweets* relacionados ao ERE, distribuídos em intervalos de tempo que categorizaram expectativas anteriores, durante e posteriores a esta estratégia de ensino adotada. Fez-se o uso de bibliotecas de código aberto para mineração de opiniões e do método Multinomial Naive Bayes, a fim de construir um procedimento que classificaria os sentimentos da comunidade em cada intervalo de tempo proposto. A escolha do método deu-se após testes efetuados com outras bibliotecas e por ser adequado para o contexto de classificação textual, como está descrito em sua documentação. A análise de dados referentes aos resultados obtidos não busca identificar soluções psicológicas, mas promover uma reflexão dos impactos do ERE nos Institutos Federais, servindo como uma pré-análise deste momento significativo para o cenário educacional da Rede Federal de Ensino Profissional e Tecnológico do Brasil.

O artigo está organizado da seguinte forma. Na seção 1 foi realizada uma introdução ao problema de pesquisa, sua relevância e a proposta de solução adotada. A fundamentação teórica e revisão bibliográfica é apresentada na seção 2. A metodologia empregada, bem como a descrição detalhada dos passos do desenvolvimento é descrita na seção 3. Na seção 4 são discutidos os resultados alcançados, e por fim, na seção 5 é apresentada a conclusão do trabalho.

2. FUNDAMENTAÇÃO TEÓRICA

A popularização das redes sociais como *Twitter*, *Facebook* e *LinkedIn* proporcionou ao usuário uma nova maneira de expressar sua opinião. A análise de sentimento textual é um processo computacional que identifica e categoriza uma opinião em um texto sem limitações de tamanho e formato que expressa uma atitude positiva, negativa ou neutra de um usuário para um determinado produto, evento ou personalidade [1]. Podendo ser extremamente útil em diversas situações, mas não tão simples de ser aplicada, principalmente em virtude da complexidade envolvida na linguagem humana. Os humanos podem interpretar facilmente declarações, entretanto pode ser difícil para a máquina en-

tender.

Muitos trabalhos referentes à análise de sentimentos têm sido feitos atualmente. Uma vez que o consumo de internet cresceu gradativamente durante a pandemia, a quantidade de interações também aumentou, e com isso as redes sociais passaram por um crescimento considerável de acesso e interações entre usuários. Uma das redes sociais mais famosas, o *Twitter*, foi uma das que mais cresceu. Segundo o Statista³, em 2021 a rede social contava com 14,5 milhões de usuários somente no Brasil, e estipula-se o acréscimo de 1 milhão de usuários em 4 anos. Para [10] o *Twitter* é um ótimo local para se obter opinião sobre determinado assunto, já que é uma rede social focada em opiniões de usuários. O comportamento dos usuários e suas postagens, variam de acordo com fatos temporais, como o lançamento de um novo smartphone, eleições presidenciais, especulações sobre a vida de artistas, etc. Publicações realizadas no *Twitter* se referem a acontecimentos atuais, sendo assim, uma excelente ferramenta para a coleta de dados em tempo real.

Diversos termos são comumente utilizados no contexto de análise de sentimentos. Um deles é a polaridade, definida por [3] como o grau de positividade ou negatividade de um texto, não necessariamente sendo uma binariedade, podendo ser tratado com uma ternariedade entre, positivo, negativo e neutro. A força do sentimento também pode ser levada em consideração durante a análise, podendo ser medida entre $-\infty$ e $+\infty$, definindo assim o quão forte é o sentimento naquele dado.

O pré-processamento de texto se trata da primeira etapa em relação à análise de sentimentos a partir de dados coletados no *Twitter*. Conforme [6], tal etapa consiste na limpeza de ruídos presentes nos dados coletados, removendo assim do HTML, *tags*, *scripts* e anúncios. Além disso, nem todas as palavras presentes no texto coletado são importantes, muitas vezes se tratam de caracteres especiais, símbolos ou palavras sem relação com o propósito da análise. A remoção desses ruídos simplifica a definição do sentimento, fazendo com que o algoritmo aplicado aja diretamente no texto sem perder tempo de execução ou evitando códigos desnecessários.

Com o dado pré-processado é necessário sua simplificação. A abordagem de [2] em sua mineração de *tweets*, trata de reduzir ao máximo os dados, porém de forma mais específica, sem a remoção generalizada de ruído como no pré-processamento. Em sua pesquisa, o *Stemming*, que é a redução de palavras flexionadas ou derivadas para uma versão mais simples, vem antes da remoção de *Stopwords*, que são palavras irrelevantes para o resultado final da frase, como por exemplo “a”, “o”, “e”.

O trabalho de [7] aplica uma técnica semelhante à presente pesquisa, realizando a coleta dos dados por meio da API *Twitter4J*. Com um período de tempo especificado entre 30 de maio a 13 de julho de 2014, buscou-se coletar os *tweets* relacionados ao jogo de futebol da seleção brasileira contra a seleção alemã. Os textos foram classificados em três polaridades, negativo, neutro e positivo, porém, realizados de forma manual.

[9] aborda de forma ampla em sua pesquisa a classificação dos *tweets*, passados por um filtro de expressões da língua portuguesa, buscando demonstrações textuais explícitas de sentimentos. Adicionalmente os autores utilizaram tratamento de localização e condição climática, para um entendi-

²<https://www.sinditelebrasil.org.br/>

³<https://www.statista.com/forecasts/1146589/twitter-users-in-brazil>

mento de como tais variáveis afetam os sentimentos dos dados.

3. METODOLOGIA

Para o desenvolvimento deste trabalho foi utilizada a linguagem de programação Python, versão 3.6. Inicialmente é feita a coleta dos *tweets* e em seguida o pré-processamento das informações. Para a categorização dos *tweets* é utilizado um método de aprendizagem de máquina, e posteriormente realizada a análise dos dados. As etapas do desenvolvimento são ilustradas na Figura 1.



Figura 1: Fluxograma do processo algorítmico em relação aos dados.

Fonte: Autores.

Para alcançar o propósito das etapas mencionadas na Figura 1, descrevemos a seguir os procedimentos referentes à obtenção e classificação da base de dados, bem como o tratamento da informação nesta fase, além da base de treinamento para a categorização dos *tweets* identificando os sentimentos expressos pelas opiniões da comunidade dos Institutos Federais.

3.1 Base de Dados

O intervalo de tempo considerado para a análise de sentimentos durante o ERE, abrangeu o período de junho de 2020 à outubro de 2021. Foram realizadas cinco coletas de dados caracterizadas conforme a Tabela 1. Os períodos foram estabelecidos de acordo com documentos oficiais do MEC, mais especificamente a Portaria nº 343 de 2020, e o Parecer CNE/CP 19/2020⁴, aprovado no dia 6 de outubro de 2020, onde foi alterada a Medida Provisória 934, que estabeleceu as normas para o ano letivo nos níveis da Educação Básica e da Educação Superior.

Tabela 1: Caracterização das coletas de dados.

Coleta	Intervalo	Significado	Tweets
1 ^a	06/20 a 08/20	Expectativa para o ERE	1679
2 ^a	09/20 a 11/20	1 ^a Aval. ERE	698
3 ^a	02/21 a 04/21	2 ^a Aval. ERE	610
4 ^a	05/21 a 07/21	1 ^a Expec. para presencial	584
5 ^a	08/21 a 10/21	2 ^a Expec. para presencial	538

As cinco coletas referem-se a distintos períodos temporais, os quais foram definidos com o intuito de compreender o comportamento e a diferença dos dados em relação a cada momento do ERE. A primeira coleta busca identificar a polaridade do sentimento em relação a expectativa do início do ensino remoto. Na segunda coleta, já com o ERE em andamento, tem-se o primeiro *feedback* efetivo sobre esta estratégia de ensino. No período de fevereiro a abril de 2021 foi considerada a ocorrência da segunda avaliação em relação ao ensino remoto, buscando verificar se após um período efetivo do ERE manteriam-se as mesmas polaridades. A quarta coleta representa temporalmente a expectativa em relação à

⁴<http://portal.mec.gov.br/docman/dezembro-2020-pdf/167131-pec019-20/file>

volta do ensino presencial, ainda em um cenário de incerteza sobre a possibilidade de retorno tendo em vista a situação da pandemia. A quinta coleta busca identificar a polaridade do sentimento, a partir da proximidade ao retorno presencial. Mediante os diferentes períodos propostos buscamos compreender a dinâmica dos sentimentos dos alunos, professores e demais envolvidos no ensino remoto.

Para a coleta dos *tweets* utilizou a biblioteca aberta *SNS-crape* que por meio de filtros de palavras sem a preocupação de ser sensível a letras maiúsculas e minúsculas, além de utilizar um filtro de linguagem para evitar buscas em outros idiomas. Nesta coleta foram especificados filtros que atingissem diversas possibilidades de conteúdo em relação aos Institutos Federais, como: “IFMG, IF, IF Ensino Remoto, IF Sabará”. A base de dados é tratada, ou seja, ocorre um pré-processamento do texto, conforme descrito na seção 2. Nessa etapa são removidos acentuações, *links*, caracteres especiais, e *emojicons* que poderiam causar dualidades nos resultados da categorização e análise dos *tweets*.

3.2 Base de Treino

Após a coleta e limpeza de dados, inicia-se o processo de treinamento e teste do modelo para categorizar os sentimentos das expressões (dados) coletadas. Considerando um aprendizado supervisionado, foram utilizados 8.199 *tweets*⁵ já rotulados entre negativo (2446 *tweets*), neutro (2453 *tweets*) e positivo (3300 *tweets*), para construir um dicionário a ser utilizado no *corpus*. A base de dados utilizada foi disponibilizada por “Minerando Dados”⁶, e conta com milhares de *tweets* relacionados à educação e a política do Estado de Minas Gerais no ano de 2017. No processo de avaliação da classificação de dados por aprendizagem de máquina, mediante o uso de modelos estatísticos é usual dividirmos nossa amostra em dados de treinamento e de teste, a fim de evitar o sobreajuste (*overfitting*), ou seja, o modelo se ajusta ao conjunto de dados, mas se mostra ineficaz em prever novos resultados.

Os dados de treinamento são utilizados para determinar os parâmetros do modelo e em seguida comparamos as predições do modelo para os dados de teste com os dados que realmente observamos, e usamos esta comparação para aferir sobre a precisão (acurácia) do nosso modelo. É necessário definir a quantidade de dados que será utilizada para o treinamento e a quantidade que será utilizada para ser testada. Análises empíricas indicam que os melhores resultados são obtidos se alocarmos 30-20% dos dados originais para teste e o restante, ou seja, 70-80% para o treinamento [5].

A fim de validar o modelo utilizado e garantir maior veracidade de sua assertividade calculamos as principais métricas para a classificação de dados, conforme apresentado na Tabela 2.

4. ANÁLISE DOS RESULTADOS

Com os dados tratados a partir das coletas realizadas e após o treinamento do modelo para a classificação dos 4.109 *tweets* coletados, fez-se a categorização dos sentimentos de acordo com as coletas e seus respectivos períodos de tempo

⁵https://github.com/stacktecnologias/stack-repo/blob/master/Tweets_Mg.csv

⁶<https://web.archive.org/web/20211229192424/https://minerandodados.com.br/>

Tabela 2: Métricas do modelo.

	Negativo	Neutro	Positivo
Acurácia	97%	97%	100%
Precisão	90%	95%	100%
Sensibilidade	90%	95%	100%
F1-score	90%	95%	100%

no contexto do ERE.

4.1 Primeira Coleta

A primeira coleta é referente a polaridade do sentimento em relação a expectativa do início do ensino remoto. Por ser uma mudança brusca de rotina e um cenário social e educacional totalmente novo, era esperado que as opiniões iniciais, ou seja, referentes à **primeira coleta**, fossem classificadas como negativas, porém, a Figura 2 mostra um recebimento favorável em relação à nova metodologia de ensino. Podem ter influenciado esse resultado, o fato de ser o início da pausa do ensino presencial e a expectativa sobre o planejamento funcional do ensino.

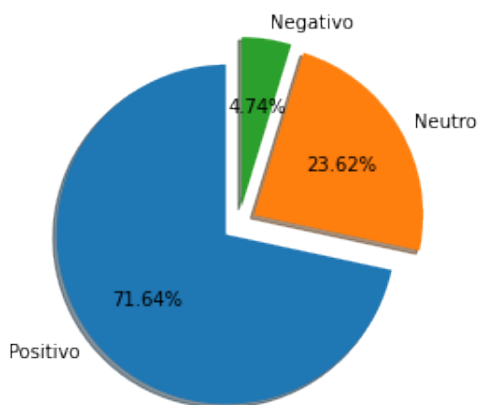


Figura 2: Gráfico de polaridades da primeira coleta.
Fonte: Autores.

4.2 Segunda Coleta

Inserido de fato no ERE, a **segunda coleta** gera uma nova possibilidade para a análise, abrindo espaço para a *feedback* do modelo em vigor. Dados do IBGE⁷ apontam um detalhe importante em relação aos alunos que foram abrangidos nos *tweets* coletados. Cerca de 49% das secretarias municipais de educação indicaram dificuldade dos estudantes em acessar a internet e a partir disso, é possível perceber uma menor interação dos alunos nas redes sociais. Apesar deste fato, percebe-se, a partir da Figura 3 a prevalência dos *tweets* “positivos” em relação aos “negativos”. Nota-se o aumento dos sentimentos “neutros” (+120%) e “negativos” ($\approx +70\%$) e o declínio dos sentimentos “positivos” ($\approx -44\%$).

⁷<https://undime.org.br/noticia/14-04-2021-13-19-segundo-ibge-43-milhoes-de-estudantes-brasileiros-entraram-na-pandemia-sem-acesso-a-internet>

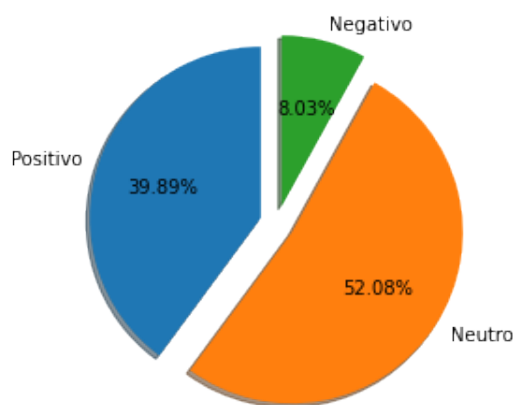


Figura 3: Gráfico de polaridades da segunda coleta.
Fonte: Autores.

4.3 Terceira Coleta

Uma questão associada à **terceira coleta** é verificar se o início do ERE poderia ter causado uma impressão diferente da que viria a ser a opinião verdadeira da comunidade, após um maior tempo em atividade do ensino remoto. Com a dominância da neutralidade na terceira coleta é possível perceber o início de um padrão de comportamento em relação às polaridades. Nota-se a partir da Figura 4 a diminuição da positividade, o aumento da negatividade e uma dominância da polaridade neutra. Tal fato pode ter sido influenciado pela grande quantidade de notícias informativas em relação aos Institutos Federais, dado a evolução sobre os protocolos de saúde e as normas de ensino. Informativos em relação a evolução do vírus geralmente não incluem nenhum tipo de opinião por se tratar de um meio de comunicação, não expressando nenhuma positividade ou negatividade, apenas apresentando fatos ou ocorridos.

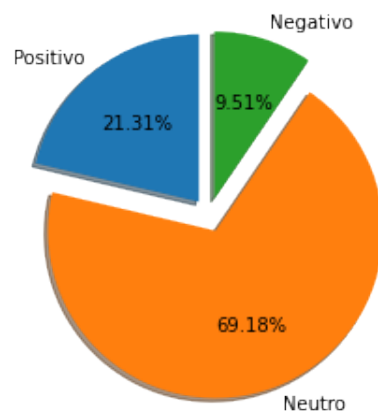


Figura 4: Gráfico de polaridades da terceira coleta.
Fonte: Autores.

4.4 Quarta Coleta

A **quarta coleta** visava captar o sentimento a respeito da expectativa de volta às aulas presenciais tendo em vista a situação da pandemia. Percebe-se, conforme a Figura

5 a ocorrência do padrão mencionado anteriormente. As diferenças percentuais da terceira para a quarta coleta são -0,42% (“positivo”), +0,76% (“negativo”) e -0,34% (“neutro”).

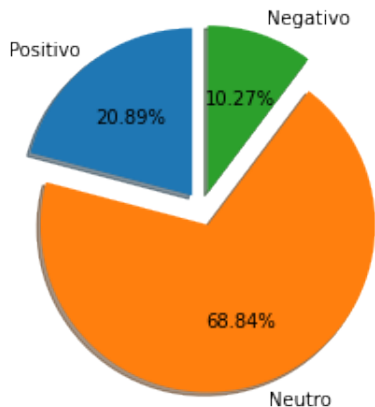


Figura 5: Gráfico de polaridades da quarta coleta.
Fonte: Autores.

4.5 Quinta Coleta

Na última coleta, com uma maior proximidade ao retorno presencial, ou seja, a **quinta coleta**, o sentimento de positividade aumenta aproximadamente 16,5% enquanto os sentimentos “neutro” e “negativo” tiveram queda, de acordo com os dados da Figura 6.

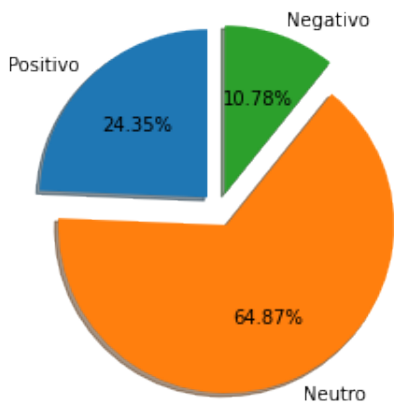


Figura 6: Gráfico de polaridades da quinta coleta.
Fonte: Autores.

A Figura 7 exibe o comportamento dos sentimentos a partir dos dados coletados nos cinco períodos de análise em torno do ERE. Verifica-se que os sentimentos “positivo” e “neutro” apresentaram maior variação, e com tendências opostas, ou seja, enquanto houve um crescimento do sentimento de neutralidade ao longo do período avaliado, constatou-se um declínio do sentimento de positividade. Nota-se que em termos absolutos, o sentimento de positividade vindo das redes sociais foi maior do que os demais sentimentos de junho de 2020 a setembro do mesmo ano, correspondendo ao período de expectativa ao ERE. No restante do período

avaliado a neutralidade prevaleceu e a partir da segunda avaliação do ERE (fevereiro a abril de 2021) os sentimentos apresentaram pouca variação, com um “leve” crescimento do sentimento de positividade a partir do segundo semestre civil de 2021.

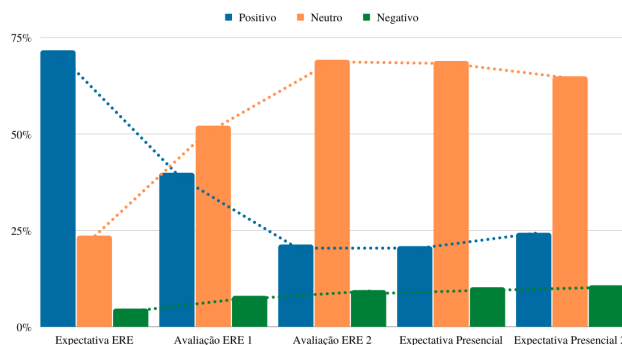


Figura 7: Evolução das coletas.

Na determinação das polaridades nota-se que o sentimento “positivo” foi bem classificado em todas as métricas. Para os demais sentimentos, o menor patamar de classificação, considerando os dados treinados, foi de 95% para a polaridade neutra e 90% para a polaridade negativa. Vale ressaltar que na acurácia, ou seja, dentre todas as classificações verdadeiras que o modelo fez, todas positivas estavam corretas, sentimentos neutros e negativos atingiram 97%. As porcentagens métricas encontradas apresentaram-se muito próximas ou exatas em 100%, ocorrido advindo de um possível *overfitting*, uma vez que os dados utilizados no treino e nos testes faziam parte de um contexto específico e correlacionado. Uma possível solução para a melhora do modelo seria o aumento da variedade de *tweets* da base de treino.

5. CONCLUSÃO

Neste trabalho propomos identificar a percepção e sentimento da comunidade relacionada aos Institutos Federais durante a vigência do Ensino Remoto Emergencial. Os dados coletados via *Twitter* foram tratados e submetidos a um modelo de treinamento utilizando aprendizagem de máquina supervisionada para classificar as opiniões coletadas como “positiva”, “negativa” ou “neutra”.

As maiores variações de sentimentos foram percebidas pelas polaridades “positiva” e “neutra”, sendo que o sentimento “positivo” apresentou um declínio até o início da segunda avaliação do ERE (fevereiro a abril de 2021) e o sentimento “neutro” uma ascensão no mesmo período. O sentimento “negativo” apresentou crescimento mais expressivo na transição da expectativa para o ERE para sua primeira avaliação.

A análise de sentimentos por meio de rede social configura uma expressiva temática de pesquisa, com diversas possibilidades de coleta e uma variedade de sentimentos a serem utilizados. O modelo de análise apresentado nesta investigação, pode ser aprimorado a ponto de interpretar com maior precisão os dados coletados, por meio de cruzamento de bases ou com a identificação de contexto para o momento da análise.

Para uma melhor performance do modelo no futuro, a inclusão de *emoticons* e símbolos no momento da análise pode agregar mais valor para a precisão dos dados. Utilizando da

substituição de um sorriso para a palavra “alegria” como um meio de simplificar a análise mantendo a ideia do símbolo.

Esperamos que estes resultados sirvam como base para uma pré-análise do sentimento gerado pelo ERE no contexto do ensino dos Institutos Federais, e que possamos aproveitar o conhecimento gerado para eventuais desafios desta grandeza no futuro.

6. AGRADECIMENTOS

Agradecemos ao Instituto Federal de Minas Gerais - campus Sabará pelo suporte na execução desse trabalho.

7. REFERÊNCIAS

- [1] S. Alam and N. Yao. The impact of preprocessing steps on the accuracy of machine learning algorithms in sentiment analysis. *Computational and Mathematical Organization Theory*, 25(3):319–335, 2019.
- [2] A. Bakliwal, P. Arora, S. Madhappan, N. Kapre, M. Singh, and V. Varma. Mining sentiments from tweets. In *WASSA@ ACL*, pages 11–18, 2012.
- [3] F. Benevenuto, F. Ribeiro, and M. Araújo. Métodos para análise de sentimentos em mídias sociais. *Sociedade Brasileira de Computação*, 2015.
- [4] M. L. Castilho and C. N. N. da Silva. A covid-19 e a educação profissional e tecnológica: um panorama das ações de acompanhamento e enfrentamento da pandemia nos institutos federais. *Revista Nova Paideia-Revista Interdisciplinar em Educação e Pesquisa*, 2(3):18–34, 2020.
- [5] A. Gholamy, V. Kreinovich, and O. Kosheleva. Why 70/30 or 80/20 relation between training and testing sets: A pedagogical explanation. Technical report, University of Texas at El Paso, 2018.
- [6] E. Haddi, X. Liu, and Y. Shi. The role of text pre-processing in sentiment analysis. *Procedia computer science*, 17:26–32, 2013.
- [7] S. M. Moraes, I. H. Manssour, and M. S. Silveira. 7x1pt: um corpus extraído do twitter para análise de sentimentos em língua portuguesa. In *Anais do X Simpósio Brasileiro de Tecnologia da Informação e da Linguagem Humana*, pages 21–25. SBC, 2015.
- [8] E. A. M. Oliveira and C. A. Silva. *Ensino Remoto Emergencial (ERE): múltiplas visões e vivências no ensino técnico e tecnológico em tempos de pandemia*, volume 1, chapter Os desafios da gestão pedagógica na vigência do Ensino Remoto Emergencial em 2020, pages 37–51. Editora Fi, Porto Alegre, 1 edition, 2022.
- [9] A. O. Rodrigues, R. C. de Melo-Minardi, and W. M. Junior. Como nos sentimos: Uma ferramenta de mineração visual de sentimentos no twitter. In *Anais do II Brazilian Workshop on Social Network Analysis and Mining*, pages 91–102. SBC, 2013.
- [10] C. A. S. Rodrigues, L. L. Vieira, L. Malagoli, and N. Timmermann. Mineração de opiniao/análise de sentimentos. trabalho acadêmico, universidade federal de santa catarina, florianópolis, 2010.